PROTEIN PURIFICATION BY USING IMMOBILIZED METAL ION AFFINITY (IMA) ADSORBENT

FIDELIA SAWAI ANAK MICHAEL MULOK

A thesis submitted in fulfillment of the
requirements for the award of the degree of
Bachelor of Chemical Engineering

Faculty of Chemical and Natural Resources Engineering
Universiti Malaysia Pahang

MAY 2009

TABLE OF CONTENTS

ABSTRACT

Effective separation and purification of proteins has been an important issue in the biomedical and pharmaceutical industries. A novel protein adsorption has been developed in biotechnology to achieve highly efficient and economical separation processes. Application in separation and purification processes often used the ability of zeolites and other molecular sieves to exclude molecules too large to enter the pores and admit smaller ones. In this study, two zeolites which are H-Y and H-Beta have been modified by adding a type of metal into each zeolite to enhance the performance of the zeolites. Three types of metals were used. They were nickel oxide, ferum oxide and zirconium oxide. The zeolite is used as an immobilized metal ion affinity stationary phase for protein purification. The adsorption of Bovine Serum Albumin (BSA) protein using modified zeolites was studied. The effect of pH on adsorption capacity was studied at three different pHs, namely 3, 5 and 8. It is found that the adsorption capacity is the highest at pH 5 which is the nearest to the pI of BSA. Increase in pH higher than the pI leads to the decrease in the adsorption capacity. This is caused by electrostatics repulsion between protein and the surface of adsorbent. Bovine serum albumin concentration was analyzed by UV/VIS Spectrophotometer. It is obvious that as the concentration is higher, the adsorption of Bovine Serum Albumin (BSA) protein is also higher. This is because as the sample is more concentrated, it contained more protein so the adsorption will also be increased. It can be concluded that the most efficient zeolite is H-Beta combined with zirconium oxide. Ismail *et al.* (2005) has said that the molecular sieve H-Beta zeolite has been explored for its ability to adsorb proteins from aqueous solution in batch experiment. Zirconium oxide is the most efficient metal compared to nickel oxide and ferum oxide. The adsorption isotherms are confirmed to be ideal to the Langmuir model.

ABSTRAK

Pengasingan dan penulenan protein yang efektif telah menjadi isu yang penting dalam industri bioperubatan dan farmasi. Penjerapan protein telah dibangunkan dalam industri bioteknologi untuk mencapai proses pengasingan yang amat efisien dan ekonomikal. Proses pengasingan dan penulenan protein mengaplikasikan kebolehan zeolite dan penapis molekul yang lain untuk menghalang molekul yang terlalu besar daripada memasuki liang-liang pada zeolite dan membenarkan molekul yang lebih kecil melaluinya. Dalam kaji selidik ini, dua jenis zeolite iaitu H-Beta dan HY telah digunakan. Zeolite-zeolite ini telah diubahsuai dengan mencampurkan sejenis logam ke dalam setiap satu zeolite supaya fungsi zeolite dapat dipertingkatkan. Tiga jenis logam telah dipilih iaitu nikel oksida, ferum oksida dan zirkonium oksida. Untuk penulenan protein, zeolite digunakan sebagai tarikan ion logam yang tidak bergerak dalam fasa pegun Kaji selidik ini telah dijalankan dengan menggunakan zeolite yang telah diubahsuai. Nilai pH yang berbeza iaitu pH3, pH5 dan pH8 telah ditetapkan bagi mengkaji kesan pH terhadap kapasiti penjerapan. Penjerapan berlaku paling tinggi pada pH5, iaitu pH yang berdekatan dengan pI protein BSA. Jika pH lebih tinggi dari pI protein, kapasiti penjerapan menjadi lebih rendah. Ini disebabkan oleh daya tolakan elektrostatik antara protein dan permukaan jerapan. Kepekatan protein BSA telah dianalisis menggunakan UV/VIS Spectrophotometer. Jika kepekatan semakin tinggi, penjerapan protein juga semakin tinggi. Ini adalah kerana lebih tinggi kepekatan, ia mengandungi lebih banyak protein jadi penjerapan menjadi lebih tinggi. Zeolite yang paling efisien ialah H-Beta bercampur dengan zirkonium oksida. Ismail *et al.* melaporkan bahawa penapis molekul seperti H-Beta telah dikaji kebolehannya untuk menjerap protein daripada larutan cecair dalam eksperimen sekumpulan. Zikronium oksida adalah logam paling efisien berbanding logam nikel dan ferum. Isoterma penjerapan bagi protein BSA ini adalah ideal dengan model Langmuir.

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

x	-	Quantity adsorbed
m	-	Mass of the adsorbent
P	-	Pressure of adsorbate
k,n	-	Empirical constants
A	-	Gas molecule
S	-	Adsorption site
θ	-	Fraction of the adsorption sites occupied
v_{mon}	-	STP volume of adsorbate
v	-	Volume
θ_E	-	Fraction of empty sites
i	-	Each one of the gases that adsorb
T	-	Temperature
ΔH	-	Entropy change
c	-	Equilibrium constant

LIST OF ABBREVIATIONS

NMR	-	Nuclear Magnetic Resonance
BSA	-	Bovine Serum Albumin
pH	-	Expressing acidity or alkalinity on a logarithmic scale
pI	-	Isoelectric point
DNA	-	Deoxyribonucleic acid
RNA	-	Ribonucleic acid
IMAC	-	Immobilized Metal Ion Affinity Chromatography
HPLC	-	High-performance liquid chromatography

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background of Study

Proteins are large organic compounds made of amino acids arranged in a linear chain and joined together by peptide bonds between the carboxyl and amino groups of adjacent amino acid residues. The sequence of amino acids in a protein is defined by a gene and encoded in the genetic code. Although this genetic code specifies 20 "standard" amino acids plus selenocysteine and - in certain archaea - pyrrolysine, the residues in a protein are sometimes chemically altered in post-translational modification: either before the protein can function in the cell, or as part of control mechanisms.

Proteins were recognized as a distinct class of biological molecules in the eighteenth century by Antoine Fourcroy and others, distinguished by the molecules' ability to coagulate or flocculate under treatments with heat or acid. For examples at the time included albumin from egg whites, blood, serum albumin, fibrin, and wheat gluten. Dutch chemist Gerhardus Johannes Mulder carried out elemental analysis of common proteins and found that nearly all proteins had the same empirical formula. The term protein to describe these molecules was proposed in 1838 by Mulder's associate Jöns Jakob Berzelius. Mulder went on to identify the products of protein degradation such as the amino acid leucine for which he found a molecular weight of 131 Da.

The difficulty in purifying proteins in large quantities made them very difficult for early protein biochemists to study. Hence, early studies focused on proteins that

could be purified in large quantities, e.g., those of blood, egg white, various toxins, and digestive/metabolic enzymes obtained from slaughterhouses.

Effective separation and purification of proteins has been an important issue in the biomedical and pharmaceutical industries. A novel protein adsorption has been developed in biotechnology to achieve highly efficient and economical separation processes. In many cases, proteins which have similar physical and chemical properties need to be separated, and thus highly selective adsorbents are desired. Microporous molecular sieves, such as zeolite Y, ZSM-5 and zeolite Beta, have played important roles in acid catalysis because of their peculiar pore structures and strong intrinsic acidities (Ismail *et al.*, 2005).

Application in separation and purification processes often used the ability of zeolites and other molecular sieves to exclude molecules too large to enter the pores and admit smaller ones. Similarly, shape-selective catalysis takes advantage of the ability of the pores to favor the admission of smaller reactant molecules, the release of the smaller reaction products molecules, or restriction of the size of transition-state complexes inside the micropores of the zeolite (Shermon, 1999).

Protein separation can be done in a number of ways. For example, using membrane chromatography, glass fiber membrane modified with short-chain organosilicon derivatives, using solution by cellulose acetate (AC)/ polycarbonate (PC) blend ultrafiltration membranes and by using immobilized metal ion affinity chromatography (IMAC).

In this study, immobilized metal ion affinity adsorbent will be used. The zeolites will be modified by inserting three different types of metal, one at the time into two different zeolites. The study will be about which one of these metals can work efficiently with each zeolite to separate the chosen protein.

Immobilized metal ion affinity chromatography (IMAC) has shown promise of isolating desired proteins from a mixture based on their difference of affinity for chelated metal ions. With its technological superiority, such as large adsorption capacity, mild separation condition, simple ligands and wide applications, IMAC has become powerful tool for biotechnological products separation, such as proteins,

amino acids and gene products. In spite of many sophisticated applications for IMAC, the theoretical analysis of immobilized metal chromatography has remained insufficient (Sun XD *et al*., 2000).

1.2 Problem Statement

The efficiency needed for protein separation in biochemical, biomedical and pharmaceutical industries has always been a concern. Many developments in biotechnology have been made in order to accomplish highly effective and economical processes. Therefore highly selective adsorbents are needed.

There is no 100% efficient ways to separate protein. When using adsorption process, one of the types of adsorbents that are used is zeolite. Pure zeolites are still not efficient enough to separate protein. Therefore, in this study, two zeolites which are H-Y and H-Beta will be modified by adding a type of metal into each zeolite to enhance the performance of the zeolites. Three types of metals will be used. They are nickel, ferum and zirconium.

1.3 Objective of the Project

Objective of the study is to use zeolite as an immobilized metal ion affinity stationary phase for protein purification.

1.4 Scope of Research Work

The scope depends on the parameters that are used in the experimental process. One of the parameters is the pH of solution which can affect the adsorption capacity. Increase in pH higher than the protein pI lead to the decrease in the adsorption

capacity for the protein. This is because of electrostatic repulsion between protein and the surface of adsorbent. Other parameters are types of metals and types of zeolites.

CHAPTER 2

LITERATURE REVIEW

2.1 Proteins

Proteins are macromolecules. They are constructed from one or more unbranched chains of amino acids; that is, they are polymers. A typical protein contains 200–300 amino acids but some are much smaller and some much larger.

Proteins have high molar, ranging from about 5000g to $1x10^7$g, and yet the percent composition by mass of the elements in proteins is remarkably constant: 50-55% of carbon; 7% of hydrogen; 23% of oxygen; 16% of nitrogen; and 1% of sulfur (Chang, 2003).

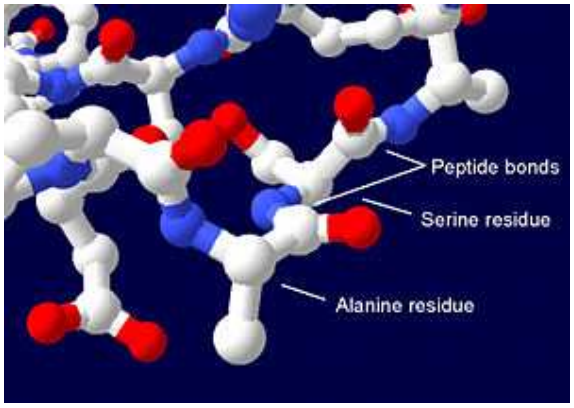

Figure 2.1: Section of a protein structure showing serine and alanine residues linked together by peptide bonds. Carbons are shown in white and hydrogens are omitted for clarity.

All amino acids possess common structural features, including an α carbon to which an amino group, a carboxyl group, and a variable side chain are bonded. Only proline differs from this basic structure as it contains an unusual ring to the N-end amine group, which forces the CO–NH amide moiety into a fixed conformation. The side chains of the standard amino acids, detailed in the list of standard amino acids, have different chemical properties that produce three-dimensional protein structure and are therefore critical to protein function. The amino acids in a polypeptide chain are linked by peptide bonds formed in a dehydration reaction. Once linked in the protein chain, an individual amino acid is called a residue, and the linked series of carbon, nitrogen, and oxygen atoms are known as the main chain or protein backbone. The peptide bond has two resonance forms that contribute some double-bond character and inhibit rotation around its axis, so that the alpha carbons are roughly coplanar. The other two dihedral angles in the peptide bond determine the local shape assumed by the protein backbone.

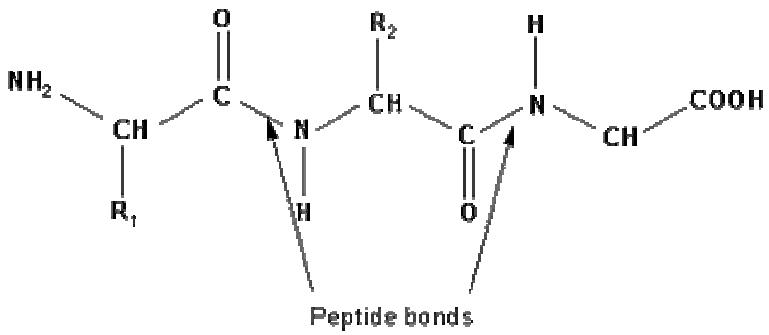

Figure 2.2: The Peptide Bond

Most proteins fold into unique 3-dimensional structures. The shape into which a protein naturally folds is known as its native state. Although many proteins can fold unassisted, simply through the chemical properties of their amino acids, others require the aid of molecular chaperones to fold into their native states.

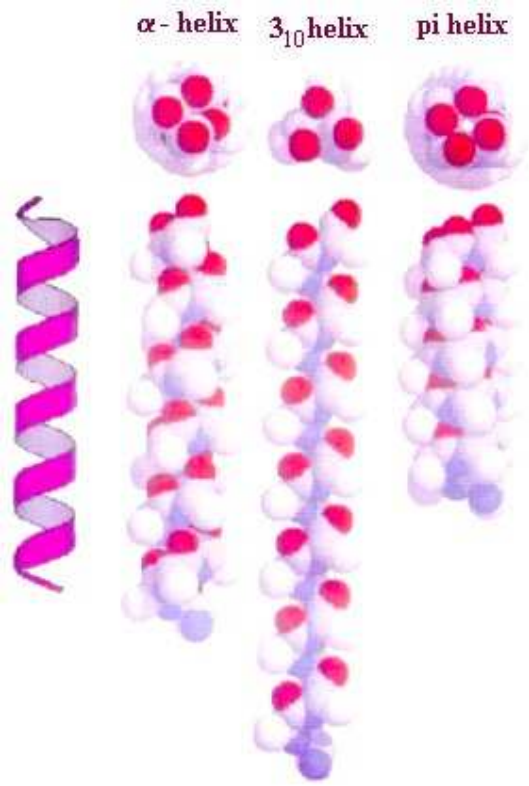

Figure 2.3: Image of alpha helix, 3_{10} helix and pi-helix

Biochemists often refer to four distinct aspects of a protein's structure: Primary structure: the amino acid sequence; Secondary structure: regularly repeating local structures stabilized by hydrogen bonds. The most common examples are the alpha helix and beta sheet. Because secondary structures are local, many regions of different secondary structure can be present in the same protein molecule; Tertiary structure: the overall shape of a single protein molecule; the spatial relationship of the secondary structures to one another. Tertiary structure is generally stabilized by nonlocal interactions, most commonly the formation of a hydrophobic core, but also through salt bridges, hydrogen bonds, disulfide bonds, and even post-translational modifications. The term "tertiary structure" is often used as synonymous with the term fold; Quaternary structure: the shape or structure that results from the interaction of more than one protein molecule, usually called protein subunits in this context, which function as part of the larger assembly or protein complex.

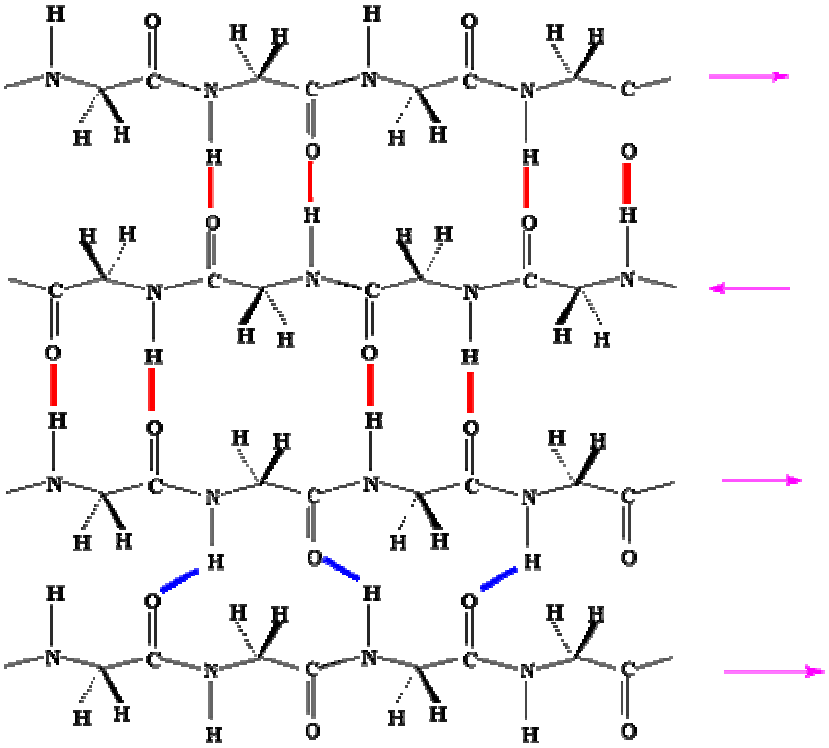

Figure 2:4: Hydrogen bond patterns in beta-sheets

Table 2.1: Parameters of regular secondary structures

Structure	Φ	ψ	n	p(Å)	A	H-bond(CO,HN)
Right-handed alpha helix [3.6_{13}helix]	-57	-47	3.6	5.4	13	i,i+2
3_{10}-helix	-74	-4	3.0	6.0	10	i,i+3
pi-helix	-57	-70	4.4	5.0	16	i,i+4
Parallel beta strand	-119	113	2.0	6.4		
Antiparallel beta strand	-139	135	2.0	6.8		

Notes: n is the number of residues per helical turn.

p is the helical pitch.

A is the atoms in H-bonded loop.

Proteins can be informally divided into three main classes, which correlate with typical tertiary structures. Firstly, fibrous proteins are composed of long linear

polypeptide chains that are bundled together to form rods or sheets. These proteins are insoluble in water and serve structural roles, giving strength and protection to tissues and cells. Secondly, globular proteins are coiled into compact shapes with hydrophilic outer surfaces that make them water soluble. Enzymes and transport proteins are globular to make them soluble in the blood and other aqueous environments in cells (Smith, 2006); and thirdly membrane proteins often serve as receptors or provide channels for polar or charged molecules to pass through the cell membrane.

Discovering the tertiary structure of a protein, or the quaternary structure of its complexes, can provide important clues about how the protein performs its function. Common experimental methods of structure determination include X-ray crystallography and NMR spectroscopy, both of which can produce information at atomic resolution. Cryoelectron microscopy is used to produce lower-resolution structural information about very large protein complexes, including assembled viruses; a variant known as electron crystallography can also produce high-resolution information in some cases, especially for two-dimensional crystals of membrane proteins. Solved structures are usually deposited in the Protein Data Bank (PDB), a freely available resource from which structural data about thousands of proteins can be obtained in the form of Cartesian coordinates for each atom in the protein.

Many more gene sequences are known than protein structures. Further, the set of solved structures is biased toward proteins that can be easily subjected to the conditions required in X-ray crystallography, one of the major structure determination methods. In particular, globular proteins are comparatively easy to crystallize in preparation for X-ray crystallography. Membrane proteins, by contrast, are difficult to crystallize and are underrepresented in the PDB. Structural genomics initiatives have attempted to remedy these deficiencies by systematically solving representative structures of major fold classes. Protein structure prediction methods attempt to provide a means of generating a plausible structure for proteins whose structures have not been experimentally determined.

Proteins are not entirely rigid molecules. In addition to these levels of structure, proteins may shift between several related structures while they perform their biological function. In the context of these functional rearrangements, these

tertiary or quaternary structures are usually referred to as "conformations", and transitions between them are called *conformational changes.* Such changes are often induced by the binding of a substrate molecule to an enzyme's active site, or the physical region of the protein that participates in chemical catalysis. In solution all proteins also undergo variation in structure through thermal vibration and the collision with other molecules, see the animation on the right.

The best-known role of proteins in the cell is their duty as enzymes, which catalyze chemical reactions. Enzymes are usually highly specific catalysts that accelerate only one or a few chemical reactions. Enzymes carry out most of the reactions involved in metabolism and catabolism, as well as DNA replication, DNA repair, and RNA synthesis. Some enzymes act on other proteins to add or remove chemical groups in a process known as post-translational modification. About 4,000 reactions are known to be catalyzed by enzymes. The rate acceleration conferred by enzymatic catalysis is often enormous - as much as 10^{17}-fold increase in rate over the uncatalyzed reaction in the case of orotate decarboxylase (78 million years without the enzyme, 18 milliseconds with the enzyme).

The molecules bound and acted upon by enzymes are known as substrates. Although enzymes can consist of hundreds of amino acids, it is usually only a small fraction of the residues that come in contact with the substrate, and an even smaller fraction - 3-4 residues on average - that are directly involved in catalysis. The region of the enzyme that binds the substrate and contains the catalytic residues is known as the active site. This was first suggested by Emil Fischer in 1894 that both the enzyme and the substrate must be geometrically compatible for them to bind and perform a certain task. This is referred to as the Lock and Key Theory.

Proteins also have structural or mechanical functions, such as actin and myosin in muscle and the proteins in the cytoskeleton, which form a system of scaffolding that maintains cell shape. Other proteins are important in cell signaling, immune responses, cell adhesion, and the cell cycle.

Proteins are also necessary in animals' diets, since animals cannot synthesize all the amino acids they need and must obtain essential amino acids from food.

Through the process of digestion, animals break down ingested protein into free amino acids that are then used in metabolism.

Type of protein that is used in this project is Bovine Serum Albumin (BSA). It is a serum albumin protein that has numerous biochemical applications including Enzyme-Linked Immunosorbent Assay (ELISAs), blots, and immunohistochemistry. Alternative uses are as a nutrient in cell and microbial culture. In restriction digests, BSA is used to stabilise some enzymes during digestion of deoxyribonucleic acid (DNA) and to avoid adhesion of enzyme to reaction tubes and other vessels. It does not affect other enzymes that do not need it for stabilisation. BSA is used because of its stability, its lack of effect in many biochemical reactions, and its low cost since large quantities of it can be readily purified from bovine blood, a byproduct of the cattle industry.

2.1.1 Bovine Serum Albumin (BSA)

2.1.1.1 Structure

The molecular weight of BSA has frequently been cited as 66,120 or 66,267, but it was revised in 1990 to 66,430. All three values are based on amino acid sequence information available at the time of publication.

BSA is a single polypeptide chain consisting of about 583 amino acid residues and no carbohydrates. At pH 5-7 it contains 17 intrachain disulfide bridges and 1 sulfhydryl group.

2.1.1.2 Physical Properties

The pI in Water at 25^0C is 5.2 and the extinction coefficient is 0.667 ml mg^{-1} cm^{-1}.The pH of 1% Solution is 5.2 to 7.Optical rotation for $[\alpha]_{259}$ is 61°and for $[\alpha]_{264}$ is

63°. The stokes radius (r_s) is 3.48 nm. The Sedimentation constant, $S_{20,W}$ X 10^{13} is 4.5 (monomer), 6.7 (dimer). The Diffusion constant, $D_{20,W}$ X 10^{7} is 5.9. The partial specific volume, V_{20} is 0.733. The intrinsic viscosity, η is 0.0413. The frictional ratio, f/f_0 is 1.30. Overall dimensions, Å is 40 X 140. The refractive index increment (578 nm) X 10^{-3} is 1.90. The optical absorbance, $A^{1\ gm/L}$ at 279nm is 0.667. The mean residue rotation, $[m']_{233}$ is 8443 and the Mean residue ellipticity is 21.1 for $[\theta]_{209}$ nm and 20.1 for $[\theta]_{222}$ nm. The estimated α-helix, in percentage (%) is 54 and estimated β-form, in percentage (%) is 18.

2.1.1.3 Solubility/Solution Stability

Albumins are readily soluble in water and can only be precipitated by high concentrations of neutral salts such as ammonium sulfate. The solution stability of BSA is very good (especially if the solutions are stored as frozen aliquots). In fact, albumins are frequently used as stabilizers for other solubilized proteins (e.g., labile enzymes). However, albumin is readily coagulated by heat. When heated to 50°C or above, albumin quite rapidly forms hydrophobic aggregates which do not revert to monomers upon cooling. At somewhat lower temperatures aggregation is also expected to occur, but at relatively slower rates.

2.1.1.4 Product description/usage

Albumins are a group of acidic proteins which occur plentifully in the body fluids and tissues of mammals and in some plant seeds. Unlike globulins, albumins have comparatively low molecular weights, are soluble in water, are easily crystallized, and contain an excess of acidic amino acids. Serum and plasma albumin is carbohydrate-free and comprises 55-62% of the protein present.

Albumin binds water, Ca^{2+}, Na^{+}, and K+. Due to a hydrophobic cleft, albumin binds fatty acids, bilirubin, hormones and drugs. The main biological function of